DOCUMENT RESUME

ED 121 727 SP 009 987

AUTHOR Denton, Jon J.
TITLE An Evaluation Model for Competency Based Teacher Preparatory Programs.
PUB DATE [76]
NOTE 15p.

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage
DESCRIPTORS Course Evaluation; *Educational Accountability; *Formative Evaluation; *Models; *Performance Based Teacher Education; Program Effectiveness; *Program Evaluation; Systems Analysis

ABSTRACT

This discussion describes an evaluation model designed to complement a curriculum development project, the primary goal of which is to structure a performance based program for preservice teachers. Data collected from the implementation of this four-phase model can be used to make decisions for developing and changing performance objectives and program components. This model provides two classes of information: data on program operations and data on student achievement. The initial component (phase one) of the evaluation model serves two assessment functions: student achievement and program effectiveness. This information is obtained from a formative test with objective-type test items. The achievement and perception data collected in phase one provide information for short term maintenance decisions for the coursework preceding student teaching. Phase two involves the confirmation of the stability of the assessment instruments. Two measurement processes constitute this phase: final cognitive assessment and observation of classroom teaching skills. Phase three centers on assessing the ability to apply teaching skills emphasized in the course work preceding student teaching. Phase IV of the model expands the effort to collect "consequence of instruction" information. Graduates of the program are asked to evaluate their preparation after one semester of classroom teaching experience on a brief questionnaire. (The report also includes a mathematical decision model, used when multiple sets of data are collected in determining whether the major goals for evaluation programs are being attained.) (DMT)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED121727

An Evaluation Model for Competency Based Teacher Preparatory Programs

Jon J. Denton
Department of Educational Curriculum and Instruction
Texas A&M University

There is a genuine need for multiple data sets to be collected at intervals in a teacher preparation program for the purpose of program evaluation. This paper discusses a four phase model which provides a structure for programmatic assessment and subsequent evaluation utilizing mathematical decision equations.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

SP009 981

Running Head: An Evaluation Model

What is a good teacher preparation program? Given the dearth of universal principles and laws of teaching, the resolution of this question may depend on the development of operational definitions of a meritorious plan, program and consequences of instruction for teacher preparation. Definitions of these elements should include: reference to the validity of the instructional objectives and the credibility of their criterion levels, the utility of instructional strategies to facilitate learning, and the effects of program management on the instructional program. An evaluation model can be implemented which synthesizes these elements into an operational plan to determine the "goodness" of a preparation program.

Moreover, literature on teacher effectiveness research (Rosenshine, B., Furst, N., 1971, 1973; Brophy, 1974; Brophy & Good, 1974; Glass, 1974; Good, Biddle, Brophy, 1975) an evaluation design (Stufflebeum, Foley, Gephart, Guba, Hammond, Merriman Provas, 1971; Schalock, 1974; Popham, 1975) provide guidelines for developing such an evaluation model. The ensuing discussion describes a model designed to complement a curriculum development project, the primary goal of which is to structure a performance based program for preservice teachers. Data collected from the implementation of this model can be used to make decisions for developing and changing performance objectives and program components. In general, using this model provides two general classes of information from which program decisions can be rendered, that is, data on program operations, and more significantly, data on student achievement. An illustration of the model for evaluating the plan, program, and consequences of instruction is presented in figure 1.

---

insert figure 1

Phase 1 - Monitoring Content Acquisition

The initial component of the evaluation model serves two assessment functions, namely, student achievement and program effectiveness. Student achievement of performance objectives for the methods and skills coursework which precede student teaching, is customarily accomplished by answering objective type test items on a formative test. Achievement data obtained from these tests provide information on individual progress in terms of the number of objectives achieved. If the instructional system allows remediation of unattained objectives, then the competencies that require remediation and retesting can be duly noted. This information provides diagnostic data for an individual's learning program and concomitant evaluation data for an instructional unit's revision. As an example of the second function of the data, if an objective in an instructional unit on classroom management requires remediation by thirty percent of the students in a class section, a difficulty apparently exists either with the objective, the instruction, the assessment, or combinations thereof.

Another procedure inherent in phase 1 requests students to complete an "Affective Antenna" after completing each unit. Attitudes of preservice teaching candidates concerning the efficacy of instructional materials and activities provide an additional dimension for determining plan and program quality.

---

insert figure 2

This attitude scale provides perceptions which complement the achievement data specific to each instructional unit. To illustrate the utility of this scale, let us assume the class summary of responses with respect to the classroom

management unit contained the following values.*

	$\bar{X}$	S.D.
Item 6 (clarity of performance objectives)	4.20	1.10
Item 8 (organization of writtem materials)	3.44	1.04
Item 9 (reasonableness of activities)	3.49	1.26

Recalling that 30 percent of the class remediated and retested one objective in this unit, we now have some evidence to suggest sources of difficulty. Given the summaries to items 6, 7, and 8 we may consider the instructional activities and related printed materials to be the "problem" in this unit.

The achievement data and perception data collected in phase 1 provide information for making short-term or maintenance decisions for the coursework preceding student teaching. Inherent in decisions rendered from these data are assumptions about the validity of the performance objectives, instructional activities, and assessment instruments of the total program. While phase 1 provides a tenuous means to check the validity of the objectives and instructional activities, confirmation of the stability of the assessment instruments is left to phase 2 of the evaluation model.

Phase 2 - Knowledge and Skill Assessment

Two different measurement processes, final cognitive assessment and observation of classroom teaching skills, constitute this assessment phase. Program measurement instruments (final tests) consisting of test items related to each instructional unit may be administered at the conclusion of the courses preceding student teaching. The purposes for administering these instruments are to assess the candidates' knowledge base of instructional techniques and

*A 5 point scoring scale with "good" ratings being designated "5" and poor ratings "1" was used to provide numerical values for each item. Responses from 59 students were summarized in this example.

to provide data for conducting item-analysis procedures on new items for future formative tests. By conducting an item-analysis on these final tests, reasonably sound formative tests can be developed. Thus, the stability of the assessment instruments left wanting in phase 1 is addressed here.

In addition, these final tests may provide another indicator of an instructional unit's quality. Because these instruments are administered at the conclusion of each course, it is assumed that responses taken collectively for each unit reflect a measure of instructional effectiveness in terms of student retention. Thus final test scores, like attitude scores and frequency counts of objectives remediated, may constitute viable sources of information for evaluating the effectiveness of the instructional plan and program.

An additional measurement function served by this phase of the model is the assessment of candidate skill performances with respect to observable teaching behaviors. Emphasis is placed on monitoring and analyzing instructional behaviors exhibited by candidates while leading learning experiences. A number of classroom observation systems, e.g., question-answer-feedback sequences, use of time, task analysis, verbatim analysis, and teacher location have been developed by various investigators (Good and Brophy, 1973; Armstrong, 1975). These systems are appropriate for both observing the frequency of occurrence and rating the quality of performance of teaching behaviors commonly found in today's classrooms. Moreover, these systems are functional in both simulated (microteaching episodes) and actual classroom settings.

Currently, data gleaned from these instruments are used for progress reports on a candidate's skill development and serve a minor role for evaluating the preparation program. As the field based preparation program expands into earlier and more intensive classroom experiences, information obtained from observation systems of this nature will assume the program assessment functions served by the formative tests described in phase 1.

Phase 3 - Student Teaching Assessment

What teaching behaviors and skills are necessary for all teachers, given the diversity of instructional materials, learning environment, and students? Rather than becoming embroiled in controversy with this question, phase 3 centers on assessing the ability of candidates to apply teaching skills emphasized in the coursework that precedes student teaching. The assumption behind this approach is: Given the dynamic nature of the classroom, different critical teaching behaviors and skills will be identified by the consortium of professionals (classroom teachers, professional organization representatives, administrators, university faculty) participating in the field experience. Recommendations from this group will then affect subsequent modifications of the program. Because of this assumption, it is important that valid and accurate assessment of candidate mastery of existing program competencies be accomplished. Rating scales which permit the classroom supervising teacher and university supervisor to rate and record a student teacher's performance are commonly used for this purpose. These rating scales should contain all competencies that must be demonstrated during the student teaching experience. When all requirements are listed, the rating scale serves as an administrative checklist to apprise supervisors and the student teacher of the competencies that remain to be achieved.

Although it is not a common practice, feedback from student teachers and their pupils can be obtained at the conclusion of the student teaching experience. Feedback of this nature can be accomplished by the administration of inventories addressed to the general classroom environment created by the student teacher's instruction. These scales provide data to corroborate supervisor ratings with both self-report information and pupil perceptions of the student teaching experience.

As with the data collected in the previous phases, responses from these various rating scales serve dual functions, namely, the final evaluation of a

a candidate's field experience and program assessment. In terms of program assessment, this phase is unique since it addresses the "consequences of instruction" issue. Although these data are derived from perceptions, qualitative values on the "goodness" of the teaching candidate's performance as it affects pupil learning are obtained. The effort to collect "consequence of instruction" information is expanded in the fourth phase of this model.

Phase 4 - Follow-up Assessment

Graduates of the teacher preparation program are requested to evaluate their preparation after one semester of classroom teaching experience on a brief questionnaire. Individuals are requested to rate both the importance of the skills emphasized in the preparation program, and the effectiveness of instruction provided in presenting these skills. Their perspectives, tempered by the realities of the classroom are meaningful sources of data for evaluating the "consequences of instruction" of the program. Seeking feedback after one semester of teaching experience, rather than at the end of the first or second year is recommended for two reasons. First, the delay in compiling the data for evaluating the instructional components at one or two year periods would practically negate the possibility of using these data for revising instructional components. Second, former students' recall of specific instructional units would likely be more accurate after a few months rather than after one or two years.

In addition to the program evaluation by the first year teacher, a parallel questionnaire is submitted to the immediate supervisor (the principal or department head) of the first year teacher. On this questionnaire the supervisor is requested to rate the importance of the skills emphasized in the teacher education program, and rate the effectiveness of the first year teacher in performing these skills.

Long term program decisions are rendered by mathematically combining these data with those collected in the three preceding phases of the model. How this is accomplished is the subject of the following discussion and example.

Mathematical Decision Model

Typically, the major goals for evaluation programs on teacher education include: (1) determining the quality of the product, that is, how well the graduate of the program performs; and (2) determining the quality of the process, or whether the program provides the appropriate learning experiences for entry into teaching. The substance of this section addresses an evaluation issue that results when multiple sets of data are collected during the course of a student's preparation to determine whether the aforementioned major goals are being attained.

A substantive issue for program decision-making is that of "weighting the data." Collected program data are expected to be additive; this can be assured if instruments used yield interval scale data and the data are collected on the same subjects over an extended period of time. It should be noted however, that the value of each type of data may not necessarily be equivalent. For example, which of the following data summaries is more important? During the instructional program, students rated a classroom management instructional unit high on an attitude rating scale. However, follow-up ratings from these students during the first year of teaching, indicated an inadequate background in classroom management techniques on similar instrumentation. These conflicting reports accentuate the problem of deciding what information is most important. For the program objective under consideration, one solution is simply to assign an "importance weight" (Iw) based on the intuition of the development staff. Follow-up data from former students might be rated Iw = .5, for example, and

student attitude data, Iw = .2. Other approaches are to perform multiple regression techniques on the data collected, and assign "importance weights" to each data component from the adjusted partial sums of squares for each variable; or determine correlation coefficients between the classification variables and use these values for the "importance weights." (Cooley, Lohnes, 1971).

Regardless of the method used, assigning weights must be repeated for each instructional unit, because the criterion variable changes. Once the importance weights are determined, a mathematical decision equation can be implemented to analyze the data. One hypothetical equation is represented by the following:

$$\sum_{i=1}^{10} Iw\bar{X}i \geqslant \text{established value}$$

$\sum_{i=1}^{10}$ the summation sign for the ten different types of data collected during the program

Iw = importance weight

$\bar{X}i$ = mean variable related to a terminal performance objective

$\geqslant$ established value - if the calculated value does not exceed some established criterion value, for example, 80 percent of maximum value, revision of the instructional program related to a specific terminal performance objective should be made.

During the 1973-74 academic year, assessment data were collected on all instructional units in a teacher preparation program. Operations performed on data from one of these units, namely, classroom management illustrate how multiple data sets may be integrated into a decision equation. Data on objectives achieved (OBACH) and objectives recycled (OBREC) were compiled from phase 1 of the model. Phase 3 yielded ratings of the classroom supervising teacher (CSUP)

and university supervisor (UNSUP), while phase 4 provided importance ratings (RGIMP) and effectiveness ratings (RGREF) from first year teachers and importance ratings (SRATIMP) and effectiveness ratings (SRATEF) from their immediate supervisors. Three of these variables were selected from the eight data sets to serve as criterion variables for different analyses. The data sets were analyzed by the RSQUARE procedure (Barr and Goodnight, 1972) to determine which predictor variables-criterion variable accounted for maximum variance. The optimal combination was determined to be the number of objectives achieved (OBACH) as the criterion variable with four predictor variables, i.e., objectives recycled (OBREC), university supervisor ratings (UNSUP), effectiveness ratings by first year teacher (RGREF) and ratings from the supervisor of first year teacher (SRATEF).

These data sets were subsequently analyzed with the REGR routine (Barr and Goodnight, 1972) to develop a prediction model. The REGR procedure applies the principle of least squares in fitting a linear model to the criterion variable and predictor variables supplied. This analysis provided adjusted partial sums of squares which in turn were employed as importance weights for the predictor variables to yield the following decision equation:

$$Y \leqslant .39\bar{X}_{UNSUP} + .09\bar{X}_{RGREG} + .23\bar{X}_{SRATEF} - .47\bar{X}_{OBREC}$$

Substituting the mean values for each variable into the equation and performing the arithmetic operations yielded 6.74. This value was then compared with the cut-off value (Y) of 7.52. (The cut-off value was determined by substituting the maximum values for the predictor variables, each multiplied by an .8 accomplishment factor, into the decision equation for the variable means and performing the appropriate arithmetic operations.) In this case, the instructional unit clearly failed to reach a "passing score" and was extensively revised.

Concluding Statements

Constructing an evaluation model to assess the plan, operation, and consequences of an educational program is both complex and challenging. Differentiating consequence information from program information, developing instrumentation that yields data compatible with information collected at a subsequent time, developing an "information weight" technique, and implementing a mathematical model for decision making are but a few of the issues that confront those embarking on an evaluation project. Regardless of the decision-making system used, developing an evaluation model provides a format for obtaining essential information for evaluating both program and product.

REFERENCES CITED

Armstrong, D.G. A Selection of Observational Techniques: Paper presented for conference on supervision of student teachers, Texas A&M University. November, 1975.

Cooley, W.W. & Lohnes, P.R. Multivariate data analysis. New York: John Wiley & Sons Inc., 1971, Chp. 3.

Glass, G.V. Teacher effectiveness. In H.J. Walberg (Ed.), Evaluating educational performance. A sourcebook of methods, instruments and examples. Berkeley: McCutchen Publishing Corp., 1974.

Good, T.L., & Brophy, J.E. Looking in classrooms. New York: Harper & Row Publishing, 1973.

Good, T., Biddle, B., and Brophy, J. Teachers make a difference. New York: Holt, Rinehart and Winston, 1975.

Popham, W.J. Educational evaluation. Englewood Cliffs, N.J.: Prentice-Hall Inc., 1975.

Rosenshine, B., Furst, N. Research on teacher performance criteria. In B.O. Smith (Ed.), Research in teacher education. A symposium. Englewood Cliffs, N.J.: Prentice-Hall, 1971.

Rosenshine, B., Furst, N. The use of direct observation to study teaching. In R. Travers (Ed.), Second handbook of research on teaching. New York:

Schalock, H.D. Notes on a model of assessment that meets the requirements of C.B.T.E. In W.R. Houston (Ed.), Exploring competency based education. Berkeley: McCutchen Publishing Corp., 1974, 209-249.

Stufflebeam, D.L., Foley, W.J., Gephart, W.J., Hammond, R.L., Merriman, H.D., Provas, M.M. Educational evaluation and decision making. Itasca, Ill: F.E. Peacock, 1971.

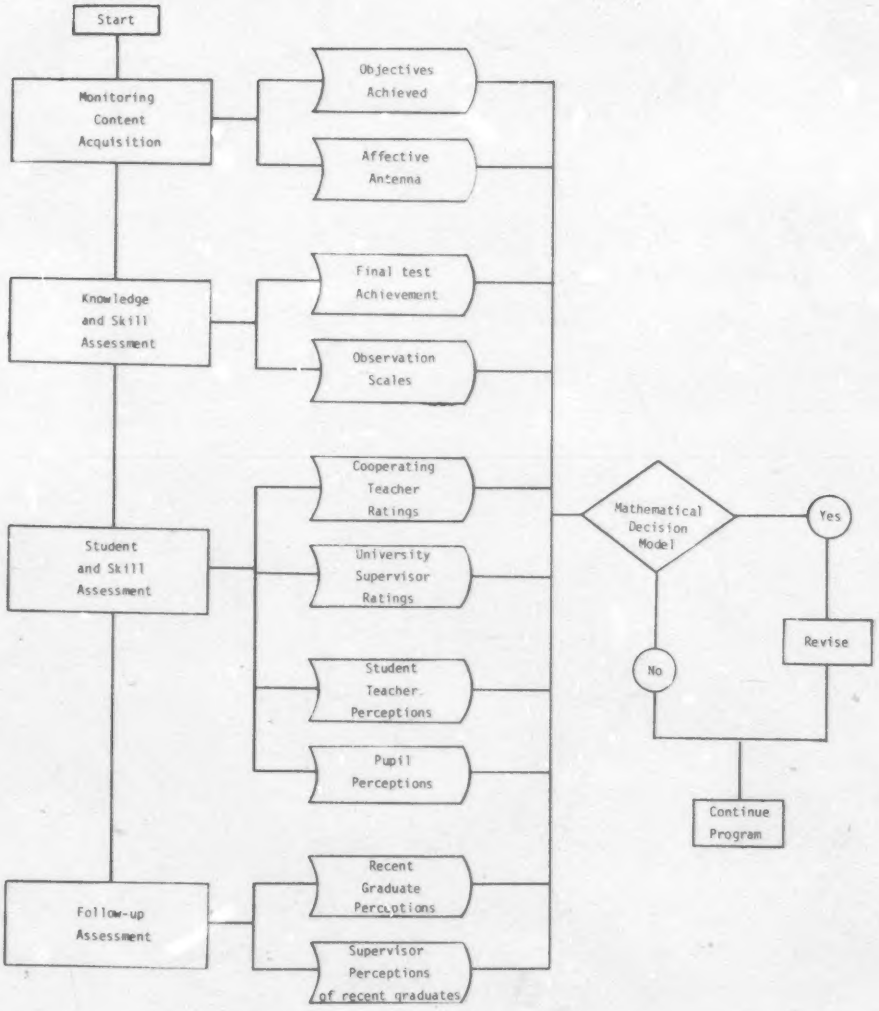

Figure 1

CBTE Evaluation Model

The following questions are being asked during this academic school year of all students taking courses in Educational Curriculum and Instruction. We want you to carefully think about the unit you have just finished in the course listed below and give your instructor your evaluation of the material this unit covered. Rate the unit, not the instructor.

Course Number _____ Section Number _____ Unit Name _____

Rate the unit on each of the following dimensions by putting an X for each dimension.

1.	ineffective	_____	_____	_____	_____	_____	effective
2.	active	_____	_____	_____	_____	_____	passive
3.	boring	_____	_____	_____	_____	_____	interesting
4.	valuable	_____	_____	_____	_____	_____	worthless
5.	organized	_____	_____	_____	_____	_____	unorganized

Now rate the unit on several other aspects by putting an X for each question.

6. To what degree were the objectives clearly stated?

Very Clear	_____	_____	_____	_____	_____	Very Unclear

7. Compared to other courses you are taking this semester, how would you rate the clarity of objectives in this unit?

Very High	_____	_____	_____	_____	_____	Very Low

8. To what degree were the unit written materials well organized?

Very Poor	_____	_____	_____	_____	_____	Very Well

9. To what degree were the unit activities reasonable (in time and effort)?

Very Reasonable	_____	_____	_____	_____	_____	Very Unreasonable

10. Overall, how would you rate the usefulness of this unit as a future teacher?

Very Low	_____	_____	_____	_____	_____	Very High

11. To what degree has the content of this unit been covered in a previous course?

Totally	_____	_____	_____	_____	_____	Not at all

Name the course(s) with this content _____

12. Does this unit need to be modified before it is used again?

Yes _____ No _____

Additional Comments:

Figure 2

Affective Antenna